

Tumor-specific retargeting of an oncogenic transcription factor chimera results in dysregulation of chromatin and transcription

Mukund Patel,^{1,2,7} Jeremy M. Simon,^{1,2,3,7} Michael D. Iglesia,² Sam B. Wu,² Andrew W. McFadden,² Jason D. Lieb,^{2,4,5} and Ian J. Davis^{1,2,5,6,8}

¹Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA; ²Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA; ³Curriculum in Bioinformatics and Computational Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA; ⁴Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA; ⁵Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA; ⁶Department of Pediatrics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA

Chromosomal translocations involving transcription factor genes have been identified in an increasingly wide range of cancers. Some translocations can create a protein “chimera” that is composed of parts from different proteins. How such chimeras cause cancer, and why they cause cancer in some cell types but not others, is not understood. One such chimera is EWS–FLI, the most frequently occurring translocation in Ewing Sarcoma, a malignant bone and soft tissue tumor of children and young adults. Using EWS–FLI and its parental transcription factor, FLI1, we created a unique experimental system to address questions regarding the genomic mechanisms by which chimeric transcription factors cause cancer. We found that in tumor cells, EWS–FLI targets regions of the genome distinct from FLI1, despite identical DNA-binding domains. In primary endothelial cells, however, EWS–FLI and FLI1 demonstrate similar targeting. To understand this mistargeting, we examined chromatin organization. Regions targeted by EWS–FLI are normally repressed and nucleosomal in primary endothelial cells. In tumor cells, however, bound regions are nucleosome depleted and harbor the chromatin signature of enhancers. We next demonstrated that through chimerism, EWS–FLI acquired the ability to alter chromatin. Expression of EWS–FLI results in nucleosome depletion at targeted sites, whereas silencing of EWS–FLI in tumor cells restored nucleosome occupancy. Thus, the EWS–FLI chimera acquired chromatin-altering activity, leading to mistargeting, chromatin disruption, and ultimately, transcriptional dysregulation.

[Supplemental material is available for this article.]

Recurrent chromosomal translocations have been associated with an increasingly wide range of human cancers. Commonly involving genes encoding transcriptional regulators, translocations can deregulate gene expression and generate structurally novel oncogenic fusion proteins (Bernheim 2010). The transforming activity of these chimeric genes typically reveals cell type specificity, suggesting that certain developmental lineages are permissive for transformation. Studies of oncogenic transcription factors have typically focused only on the fusion products or their target genes and often in heterologous cells, limiting insights into the relative influence of chimerism and cell lineage.

Ewing Sarcoma, a bone tumor of children and young adults, is characterized by translocations that fuse a member of the TET family to a member of the ETS transcription factor family (Delattre et al. 1992; Sorensen et al. 1994; Jeon et al. 1995). Identified in 80%–85% of Ewing Sarcoma, t(11;22)(q24;q12) results in an in-frame fusion of *EWSR1* to *FLI1* (Delattre et al. 1992). EWS–FLI has been shown to be a potent transcriptional modulator critical for transformation (Kinsey et al. 2006; Smith et al. 2006). Structure-function experiments have demonstrated that the EWSR1 domain contributes

transactivation activity, whereas the FLI1 domain directs DNA binding, and both are required for transformation (Lessnick et al. 1995; Jaishankar et al. 1999). EWS–FLI mediates oncogenesis by directly or indirectly regulating genes necessary for transformation. Despite evidence that EWS–FLI is necessary for transformation, ectopic expression of EWS–FLI fails to activate similar genetic programs or transform most human cell lines, indicating that cell specificity is a major determinant of EWS–FLI activity (Lessnick et al. 2002; Miyagawa et al. 2008; Riggi et al. 2008).

FLI1, a member of the ETS family, is an important developmental transcription factor (Rao et al. 1993). *Fli1* deletion in mice results in embryonic death from hemorrhage associated with aberrant hematopoiesis and vasculogenesis, supporting a role in endothelial development (Bartel et al. 2000; Spyropoulos et al. 2000; Liu et al. 2008). Translocations involving ETS members have been implicated in other cancers, including prostate adenocarcinoma (Tomlins et al. 2005). The function of EWSR1 is less well understood; however, reports suggest participation in transcription or RNA splicing (Ohno et al. 1994; Bertolotti et al. 1998). EWSR1-deficient mice die prior to weaning and show defects in B-cell development and meiosis (Li et al. 2007). Other translocations involving EWSR1 have been identified, resulting in chimeras with ATF1 and WT1 in Clear Cell Sarcoma and Desmoplastic Small Round Cell Tumors, respectively (Zucman et al. 1993; Benjamin et al. 1996; Tomlins et al. 2005).

⁷These authors contributed equally to this work.

⁸Corresponding author.

E-mail ian_davis@med.unc.edu.

Article published online before print. Article, supplemental material, and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.125666.111>.

To characterize the changes in genomic localization and transcriptional output due to chimerism, we compared EWS–FLI with FLI1 in Ewing Sarcoma and primary human endothelial cells. We integrated genomic targeting with gene expression profiling and found that in tumor cells, EWS–FLI associated with distinct genomic regions lacking canonical ETS binding sites and activating a set of genes associated with a transformed phenotype. However, in endothelial cells, genomic targeting and gene regulation were similar to that of FLI1. We then examined the influence of chromatin on this differential targeting by analyzing chromatin structure and histone modifications. We found that in Ewing cells, EWS–FLI-targeted sites exhibited features characteristic of enhancer elements and were bound by RNA polymerase II. Moreover, EWS–FLI silencing resulted in increased nucleosome occupancy of these regions. In endothelial cells, a subset of these regions, normally associated with repressive chromatin, became nucleosome depleted upon EWS–FLI expression. These data establish EWS–FLI as a pioneer factor capable of inducing and maintaining open chromatin architecture.

Results

Chimerism and cell lineage influence genomic targeting

To compare EWS–FLI with its parental protein FLI1, we developed a lentiviral delivery approach that permitted concurrent silencing of endogenous EWS–FLI or FLI1 and expression of an epitope-tagged version of either EWS–FLI or FLI1 (Fig. 1A). Lentiviral knockdown replacement was performed in a Ewing Sarcoma tumor-derived cell line (EWS502) and primary human endothelial cells (HUVEC). HUVEC were selected as they abundantly express FLI1, and FLI1 has been implicated in endothelial development (Spyropoulos et al. 2000; Puskaszeri et al. 2006; Liu et al. 2008). Genomic localization of each protein was examined by chromatin immunoprecipitation, followed by next-generation sequencing (ChIP-seq) (Supplemental Table 1). Gene expression was also examined using exon microarrays. The lentiviral knockdown-replacement strategy offered a number of experimental benefits to facilitate comparative genomic analyses. First, this approach allowed the direct comparison of each transcription factor in the same cell type as well across cell types. Second, despite heterogeneity in viral transduction, the use of lentivirus enabled the titration of protein expression, permitting approximation of endogenous protein levels while achieving efficient knockdown (Fig. 1B). Third, expression of a shRNA directed to the 3' UTR of FLI1 (able to target both endogenous EWS–FLI and FLI1, but not the transduced genes that do not contain the 3' UTR) in all experimental conditions minimized the possibility for detecting off-target effects. Finally, the use of a common and robust antibody for chromatin immunoprecipitation circumvented issues of antibody sensitivity, specificity, and antigenic variability, factors that can complicate downstream comparisons, as recently demonstrated (Egelhofer et al. 2010).

We examined cell proliferation after EWS–FLI knockdown in the presence or absence of ectopically expressed EWS–FLI or FLI1. In the tumor cells, transduced EWS–FLI, but not FLI1, rescued the growth arrest resulting from endogenous EWS–FLI silencing (Fig. 1C; Supplemental Fig. S1). These data support previous reports indicating the requirement of EWS–FLI for cell proliferation (May et al. 1993a,b; Tanaka et al. 1997; Smith et al. 2006). Silencing of FLI1 with or without ectopic expression of EWS–FLI did not affect the proliferation of endothelial cells under the conditions tested (data not shown).

Differential activities of EWS–FLI and FLI1 could result from either of two mechanisms. The transcription factors could target

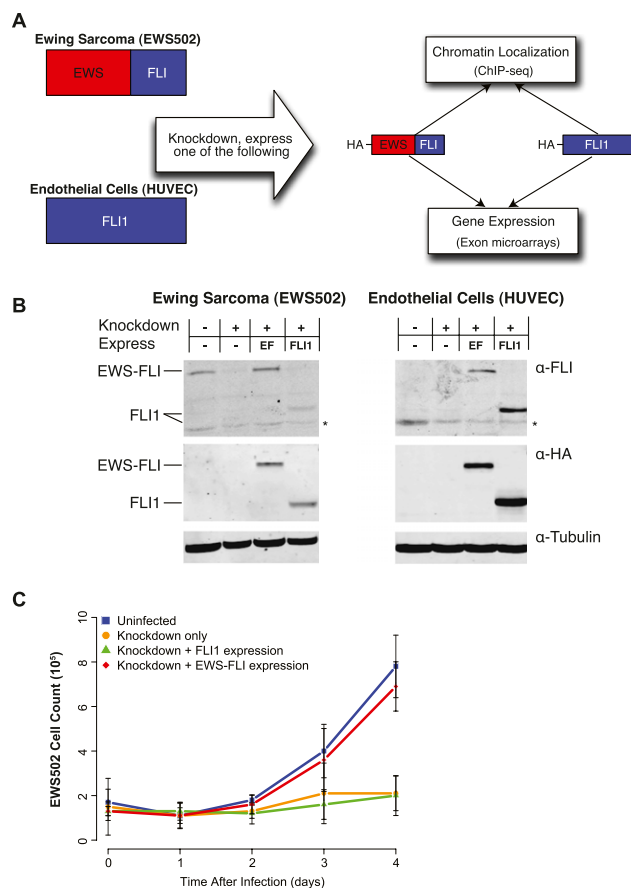


Figure 1. Experimental schema for lineage-specific transcription factor silencing and expression. (A) Ewing Sarcoma (EWS502) cells and primary human endothelial cells (HUVEC) were transduced with lentivirus expressing FLI1 3' UTR-directed shRNA and HA epitope-tagged versions of EWS–FLI or FLI1. (B) Anti-FLI1 or anti-HA immunoblots of Ewing Sarcoma cells (EWS502) or endothelial cells (HUVEC) demonstrating concurrent silencing and replacement with HA–EWS–FLI (EF) or HA–FLI1. Tubulin serves as a loading control. Asterisks indicate where a background band runs at a similar molecular weight as endogenous FLI1. (C) After EWS–FLI1 silencing alone (Knockdown) or together with ectopic EWS–FLI1 or FLI1 expression, EWS502 cells were counted. EWS–FLI expression, but not FLI1, rescues the effect of knockdown on proliferation.

similar genomic sites due to their common DNA-binding domain, but vary in their ability to modulate gene expression. Alternatively, chimerism could result in genomic retargeting and subsequent differences in transcriptional regulation. To test these two hypotheses, we performed ChIP-seq for EWS–FLI and FLI1 in both EWS502 and HUVEC. Analyzing only high-quality, uniquely aligned reads (Supplemental Table 1), binding sites for each factor were determined using the Zero Inflated Negative Binomial Algorithm (ZINBA), a flexible statistical model that adjusts for the effects of GC content, mappability, and copy-number variation (Rashid et al. 2011). We identified 7172 and 13,878 EWS–FLI binding sites in EWS502 and HUVEC, respectively. FLI1 bound 18,958 regions in EWS502 and 39,439 regions in HUVEC (Fig. 2A; Supplemental Table 2). The greater number of EWS–FLI binding sites identified in this study compared with previous ChIP-chip and ChIP-seq approaches (Gangwal et al. 2008; Guillon et al. 2009; Wei et al. 2010) likely reflects greater sequencing depth and enhanced antibody sensitivity. Since we examined targeting of transduced protein, which,

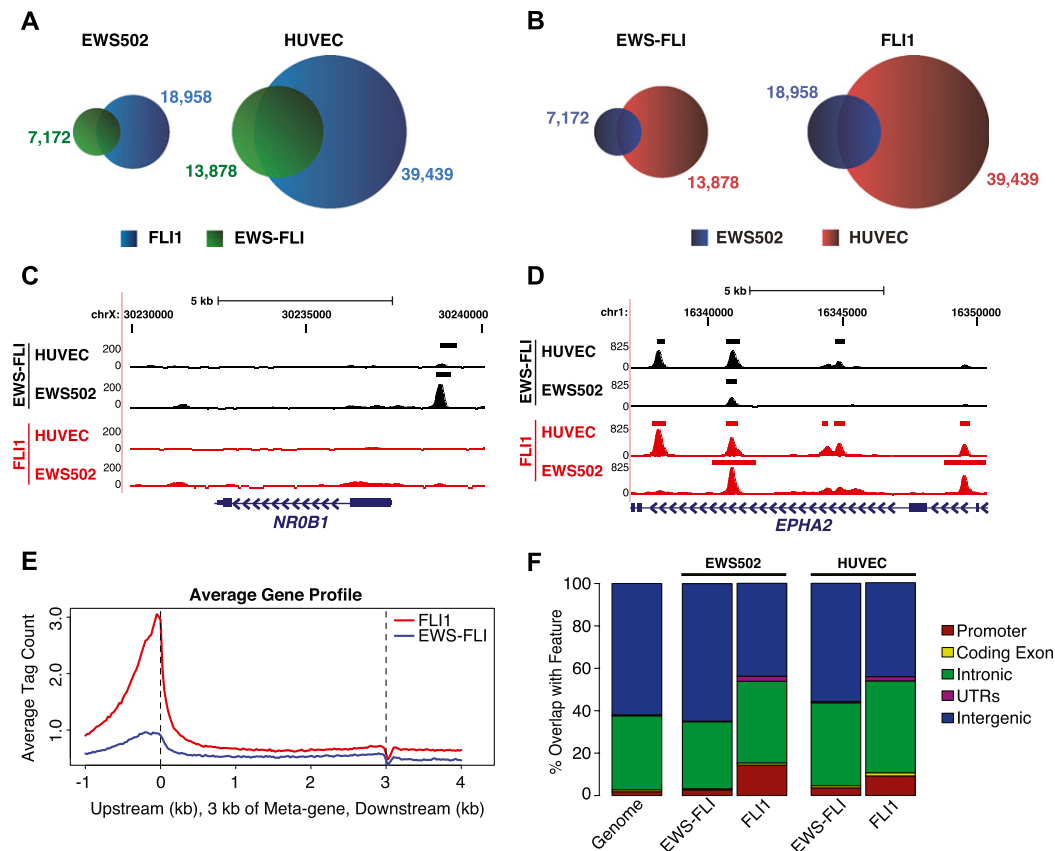


Figure 2. Chimerism alters ETS-mediated targeting. (A,B) Venn diagrams showing the number of unique and overlapping EWS-FLI and FLI1 binding regions within the same cell type (A) or across cell types (B). (C,D) UCSC Genome Browser screenshots of EWS-FLI and FLI1 ChIP-seq signal at two genes: *NROB1* (C) and *EPHA2* (D). Horizontal bars indicate targeted sites identified by ZINBA. Tag counts are shown in the y-axis. (E) Meta-gene profile of EWS-FLI and FLI1 ChIP-seq reads; 1 kb upstream of the TSS through 1 kb downstream from transcriptional termination is represented. (F) Percent overlap of ZINBA-identified EWS-FLI and FLI1 binding sites with major functional genomic features. Genomic distribution of features (Genome) is shown for comparison.

despite our efforts, varied somewhat from endogenous levels, it is possible that aberrant targeting could be detected. However, despite variation in expression and the use of different tumor cells, nearly 75% of the sites previously identified by ChIP-seq (Guillon et al. 2009) overlap the regions bound by EWS-FLI in this study.

Examination of specific genomic loci demonstrated the contribution of chimerism and cell lineage to targeting (Fig. 2A,B). For example, a site near *NROB1* previously shown to be occupied by EWS-FLI (Kinsey et al. 2006; Gangwal and Lessnick 2008; Gangwal et al. 2008; Garcia-Aragoncillo et al. 2008) was bound by EWS-FLI, but not FLI1, in both cell types (Fig. 2C; Supplemental Fig. S2A). In contrast, sites around the ephrin receptor, *EPHA2*, revealed a more complex pattern, where common as well as transcription factor- and cell type-specific sites were identified (Fig. 2D; Supplemental Fig. S2B). Overall, in the tumor cells 46% of EWS-FLI sites overlapped FLI1 sites, whereas 75% of EWS-FLI sites overlapped with FLI1 in HUVEC (Fig. 2A). Comparing targeting across cell types, only ~50% of EWS-FLI or FLI1 binding sites were shared between EWS502 and HUVEC (Fig. 2B).

We next examined the raw ChIP-seq signal over all genes to compare the association of EWS-FLI and FLI1 with transcription and genic regulatory elements (Fig. 2E). In tumor cells, the FLI1 signal was greater at transcriptional start sites (TSS), in the proximal upstream region, and through the gene body compared with EWS-FLI. Given the relative absence of EWS-FLI signal at these genic regions,

we compared the overall genomic distribution of binding sites (Fig. 2F). Again, in tumor cells FLI1 showed greater association with promoters and 5' and 3' UTRs than EWS-FLI. Compared with FLI1, EWS-FLI bound more frequently at distal intergenic regions (>60% of sites). We observed only a slight skew toward intergenic binding by EWS-FLI in endothelial cells. These data suggest that in both cancer and normal cells, FLI1 more frequently targets gene-proximal sites, and chimerism leads to retargeting to intergenic regions. However, chimerism-induced retargeting is significantly mitigated by cell lineage.

EWS-FLI and FLI1 regulate divergent gene programs

To explore the transcriptional implications of genomic retargeting, EWS-FLI and FLI1-associated gene expression changes were examined using exon microarrays in both cell types. We identified differentially regulated genes by comparing gene expression under conditions of EWS-FLI or FLI1 knockdown only to knockdown with ectopic transcription factor expression (Supplemental Table 3). Transcriptional changes may be directly or indirectly related to ectopic transcription factor expression. Although EWS-FLI occupied fewer genomic sites than FLI1, it modulated the expression of three times as many genes as FLI1 in tumor cells, though this skew was diminished in HUVEC (Supplemental Fig. S3A). Genes regulated by FLI1 or EWS-FLI were mostly distinct (40%–45% shared in either

cell type). However, of the genes commonly modulated by both factors, 97% were regulated concordantly in HUVEC, whereas opposing effects on gene expression were frequently observed in tumor cells (41% of coregulated genes) (Supplemental Fig. S3B). Cell type-specific regulation was also evident. Only 34% of genes modulated by EWS-FLI were shared across the two cell types, whereas only 12% of genes modulated by FLI1 were shared.

The classes of genes regulated by EWS-FLI and FLI1 also differed significantly (Supplemental Fig. S4). Approximately one-third of the genes modulated by EWS-FLI in tumor cells were implicated in cancer or cell cycle regulation; the identification of these categories support previous studies of gene regulation by EWS-FLI (Riggi et al. 2005, 2008; Smith et al. 2006; Kauer et al. 2009). In contrast, FLI1 expression in tumor cells induced genes associated with hematopoiesis, hematological system development and function, and cellular development, including genes of the ephrin, thrombin, and relaxin signaling pathways. In endothelial cells, similar gene ontologies were modulated by both transcription factors. These data suggest that cell type greatly influences the impact of chimerism on transcriptional output.

Differentially targeted regions are marked by DNA sequence and regulatory variation

Since EWS-FLI and FLI1 binding sites were mostly distinct, we hypothesized that additional factors might specify EWS-FLI or FLI1 targeting in a transcription factor- or cell type-specific manner. We utilized hierarchical clustering to identify binding sites that most discriminated transcription factor or cell type (Fig. 3A). Six major clusters of binding sites emerged. These clusters exhibited both transcription factor- and cell type-dependence, with cell type being the primary determinant for the majority of sites. Sites in clusters 1–3 exhibited greater binding in tumor cells, whereas those in clusters 4–6 (representing 74% of the sites) were enriched in endothelial cells. The finding of HUVEC-specific clusters containing sites bound by both EWS-FLI and FLI1 further supports that in a normal cellular environment these transcription factors share similar targeting.

Testing for associations between each cluster and gene expression demonstrated that sites in clusters 5 and 6 tended to be located near differentially expressed genes from both cell types (Fig. 3B). Approximately 15% of differentially regulated genes contained at least one of these sites within 25 kb. Furthermore, genes that contained a TSS flanked by a site in clusters 5 or 6 (± 25 kb) were significantly more likely to be modulated by EWS-FLI or FLI1 in HUVEC (Fig. 3C). Genes harboring cluster 6 sites were frequently

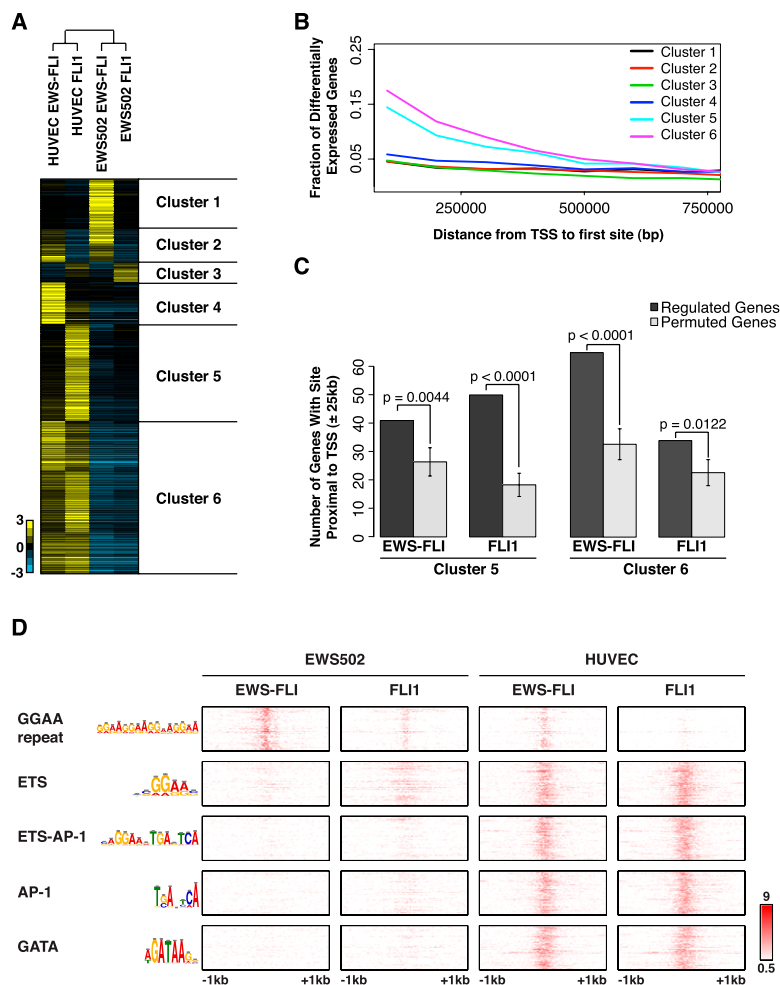


Figure 3. Hierarchical clustering identifies cell and transcription factor-specific variation in genomic targeting. (A) Hierarchical clustering of 6525 binding sites that exhibited the widest variation in signal across transcription factors or cell types. Each row was median-centered and colored based on the average read count across the region. (B) Distance (bp) from the transcriptional start site of the union set of differentially expressed genes to the closest site from clusters 1–6. (C) Number of EWS-FLI or FLI1 differentially expressed genes in HUVEC containing a cluster 5 (left) or cluster 6 (right) site within 25 kb of its TSS, compared with 10,000 permutations of all RefSeq genes. (D) Normalized log₂ ChIP-seq signal around the midpoint of identified de novo transcription factor motifs derived from the sequences underlying sites in each cluster. Clusters 1 and 2 were merged for the composite GGAA microsatellite motif (1362 rows). Clusters 3–6 were merged for ETS (682 rows), ETS-AP1 (2780 rows), AP1 (1903 rows), and GATA (812 rows). Color was assigned on a log₂ scale from 0.5 to 9.

up-regulated (82% and 88% for EWS-FLI and FLI1, respectively), however, genes proximal to cluster 5 sites lacked this skew toward up-regulation. Interestingly, these data suggest that although FLI1 targets sites in both clusters 5 and 6 equally, the potential for occupancy of EWS-FLI at these sites characterizes elements that are functionally distinct. Since the sequence composition of sites in clusters 5 and 6 were indistinguishable (see below), it is possible that chromatin differences that permit EWS-FLI binding also favor enhancer activity.

Using the Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al. 2010), we observed that regions defined by these clusters were strongly associated with specific biologically relevant ontologies independent of our gene expression data (Supplemental Fig. S5; complete list in Supplemental Table 5). In support of a direct regulatory role, EWS-FLI-specific binding sites (clusters 1 and 2) were significantly associated with genes regulated in cells engineered to express EWS-FLI. Interestingly, cluster 2 sites (common to

both HUVEC and EWS502) were associated with genes involved in mesodermal and craniofacial development, whereas sites in clusters 5 and 6 (specific for HUVEC) were largely associated with genes known to play roles in vascular development, hypoxia response, and HIF1A overexpression. These data corroborate our expression-based gene ontology observations (Supplemental Fig. S4).

To identify other features of these differentially bound regions, we performed de novo detection of over-represented motifs from each cluster (Sharov and Ko 2009) and then matched these to known motifs in TRANSFAC using TOMTOM (Matys et al. 2003). A total of 94% of sites in clusters 1 and 2, which exhibited an EWS-FLI-specific pattern in the tumor cells, contained a tetranucleotide repeat harboring the core of the motif bound by ETS family members such as FLI1 (Fig. 3D; see below). Sequence elements identified in clusters 3–6 were more highly varied but commonly contained the canonical ETS motif (see below).

Chimerism retargets EWS-FLI to tandem tetranucleotide repeats

De novo motif detection on the sequences uniquely bound by EWS-FLI in sarcoma cells (represented in clusters 1 and 2) identified

a GGAA-containing tetranucleotide microsatellite repeat. EWS-FLI binding to these sequences had been observed in recent studies (Gangwal et al. 2008; Guillon et al. 2009; Wei et al. 2010). The number of tandem repeats bound by EWS-FLI was higher than expected by chance in both cell types, although tumor cells demonstrated greater enrichment (Fig. 4A). Examination of perfect sequential repeats revealed maximum enrichment at ~14 tandem elements. Periodicity in the length of enriched repeats was observed with a preference for 8, 12, and 14 repeat units. GGAA-containing repeats, as annotated by RepeatMasker, which were bound by EWS-FLI in EWS502, were significantly longer than those not bound, with a median length of 100 bp (Fig. 4B). The difference in lengths reflects the analysis of either perfect (Fig. 4A) or imperfect (Fig. 4B) repeats. Unexpectedly, FLI1 also bound these repeats, although with much lower frequency in both cell types, suggesting that the ability to target these sites is not exclusive to EWS-FLI but rather reflects the enhancement of a native characteristic.

We directly compared the binding of EWS-FLI and other fusions characteristic of Ewing Sarcoma at sites containing either the tetranucleotide repeat or the canonical high-affinity ETS motif. The EWS-ERG and FUS-ERG fusions (Sorensen et al. 1994; Shing et al. 2003) were epitope tagged and expressed at similar levels to

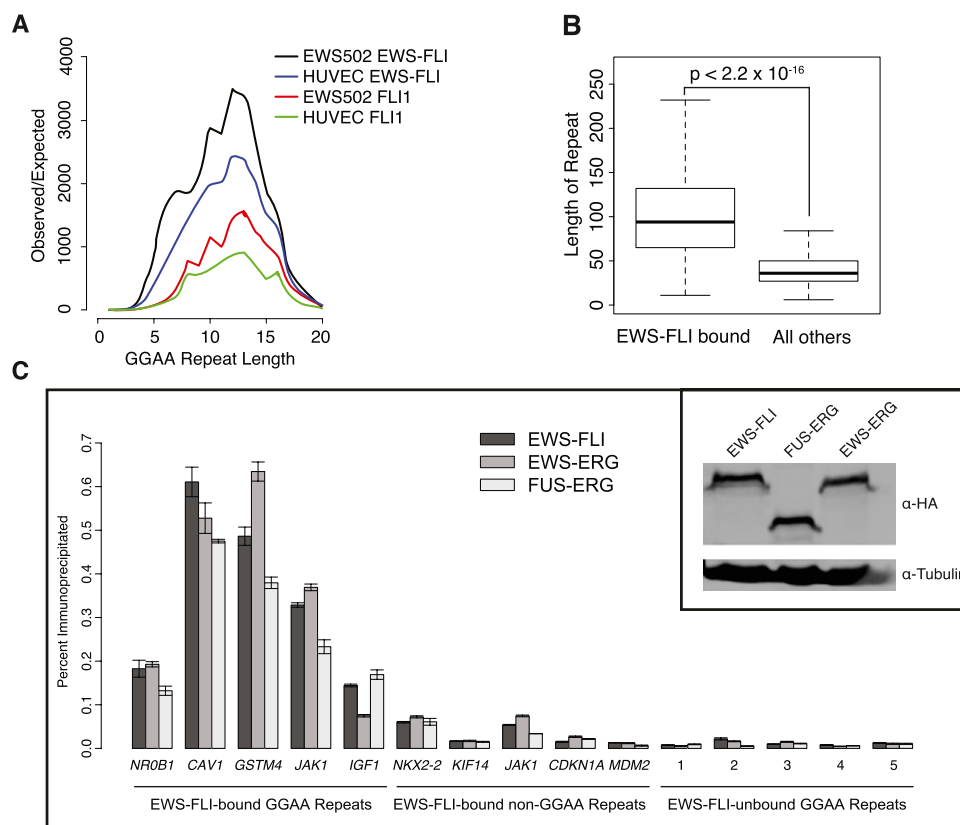


Figure 4. EWS-ETS fusions target GGAA-containing microsatellite repeats. (A) Tandem GGAA repeats identified in EWS-FLI and FLI1 binding sites in EWS502 and HUVEC were compared with those detected by 1000 permutations of the identical number of regions over the mappable genome, maintaining chromosomal distribution. All lengths exceeding one repeat were significant to $P < 0.0001$. To permit plotting lengths for which the permuted value was zero, 0.1 was added to each observed and expected value. (B) The lengths of repeat regions annotated by RepeatMasker bound by EWS-FLI in EWS502 were compared with those unbound in mappable regions of the genome. Regions bound by EWS-FLI contained significantly longer repeats as measured by a t -test. (C) ChIP-qPCR on chromatin isolated from EWS502 cells expressing the various Ewing Sarcoma fusions. Results are shown as a percent of input control. Overall, greater binding is identified at EWS-FLI-bound regions near differentially expressed genes that contained GGAA repeats (NR0B1, CAV1, GSTM4, JAK1, IGF1) compared with those that bound EWS-FLI but did not harbor a repeat (NKX2-2, KIF14, JAK1, CDKN1A, MDM2). Five control repeat-containing regions are included, and error bars represent the standard error of three replicates. (Inset) Western blot showing exogenous expression of HA-EWS-FLI, HA-FUS-ERG, and HA-EWS-ERG in EWS502 cells. Tubulin serves as a loading control.

endogenous EWS–FLI in conjunction with EWS–FLI silencing in EWS502 (Fig. 4C, inset). All fusion proteins tested demonstrated a greater enrichment at sites containing tandem repeats than canonical high-affinity ETS motifs (Fig. 4C; Supplemental Fig. S6). These data corroborate EWS–ERG ChIP (Wei et al. 2010) and support the general property of TET–ETS fusions to occupy these elements in a chromatinized genomic context. Moreover, these data suggest that repeat-containing sites are more likely to be bound than the canonical sites.

In light of recent studies suggesting a length requirement for microsatellite enhancer function (Gangwal and Lessnick 2008; Gangwal et al. 2008), we examined EWS–FLI and FLI1 sites containing five or more tandem repeats. Approximately 30% of regions uniquely bound by EWS–FLI in either cell type contained these long tandem repeats, and they were nearly undetected in regions unique to FLI1 in either cell type (0.2% and 0.04% in EWS502 and HUVEC, respectively). In agreement with the previous studies, we found that these EWS–FLI-bound GGAA repeats were more proximally located to genes up-regulated by EWS–FLI (Supplemental Fig. S7); a similar trend was observed for FLI1-bound repeats near genes induced by FLI1 (data not shown). FLI1 bound more proximally to FLI1-modulated genes compared with EWS–FLI around EWS–FLI-regulated genes (Supplemental Fig. S8; Supplemental Table 4), suggesting that in the context of chromatin, tetranucleotide repeats may function primarily as transcriptional enhancers and can be located distally from genes.

We then asked whether the preference of EWS–FLI for tetranucleotide repeats resulted from increases in repeat lengths in Ewing Sarcoma cells. Repeat-length variation could result from repeat expansion and selection during tumor development. Previous studies that examined a small number of randomly selected tetranucleotide repeats failed to demonstrate a difference between tumor cells and the reference genome (Gangwal et al. 2008; Guillon et al. 2009). We compared the lengths of several tetranucleotide repeats occupied by EWS–FLI in tumor cells with the same regions in HUVEC and the reference genome (hg18). Regions amenable to evaluation were limited due to the challenges inherent in primer design for repetitive regions. However, a region, located in an intron of *IGF1*, exhibited mono- or biallelic presence of a sequence longer than that predicted by the reference genome in four of seven Ewing Sarcoma cell lines (Supplemental Fig. S9A). Sequencing of this region from EWS502 cells confirmed that the difference resulted from nine additional repeats (Supplemental Fig. S9B). Interestingly, we observed an extremely faint band running at approximately the same molecular size as the expanded region in the pooled endothelial cells. It is thus possible that expansion (relative to the reference genome) represents the selection for an allelic variant present in the population.

Combinatorial DNA-binding motifs distinguish endothelial cell-targeted sites

Canonical ETS motifs were identified in 72% of sites in cluster 3 (largely specific to FLI1 in tumor and endothelial cells) and in sites in clusters 4–6 (specific to endothelial cells but bound by both transcription factors). The presence of ETS motifs in cluster 3 demonstrates that FLI1 retains the ability for context-dependent targeting even in sarcoma cells. Strikingly, the motif for the AP1 complex was detected at nearly the same frequency as ETS at clusters 3–6 sites. Remarkably, of the sites containing a computationally derived AP1 motif, 76% overlapped binding sites of c-Jun, a member of the AP1 complex, detected by ChIP-seq in HUVEC (The ENCODE Project Consortium 2004). In addition to isolated AP1 motifs, composite

ETS and AP1 motifs were observed at ~46% of the sites in clusters 3–6. We explored variation in the spacing between the ETS and AP1 motifs; ~25% of the sites revealed separation of 1 bp with spacing increments of 2–10 bp, each accounting for an additional 9% to 12% of sites. The composite nature of this ETS–AP1 motif is suggestive of cooperative binding at these sites. The GATA motif was also observed in ~15% of sites from clusters 3–6. The ETS motif found within these regions demonstrated context specific sequence variation. ETS sites in isolation typically contained a C preceding the GGAA core, matching the canonical ETS motif (Wei et al. 2010). However, ETS motifs in composite sites with AP1 were preceded by an A. Similar motif variation had been observed in the tandem binding sites of ETS1 with RUNX1 (Hollenhorst et al. 2009).

Since clusters 3–6 contained a similar set of motifs, we next compared the intensity, location, and specificity of binding at sites (from Fig. 3A) that contained a specific motif (Fig. 3D). Tetranucleotide repeats were preferentially and centrally bound by EWS–FLI. In the tumor cells, FLI1 demonstrated a modest ability to interact with some of these sites as previously noted (Fig. 4A). In endothelial cells, the signal intensities and positions of FLI1 and EWS–FLI were similar around ETS, ETS–AP1, AP1, and GATA motif-containing sites. In the tumor cells, FLI1 bound these sites, although with far less signal intensity, again demonstrating the tendency of FLI1 to function normally in tumor cells. Since de novo motif identification may preclude detection of less common motifs, we examined signals from HUVEC around sites containing computationally predicted motifs for MYC:MAX, NFkB, STAT, PPAR, HNF4A, and CREB (Xie et al. 2009). Although these sites represented <1% of those analyzed, similar patterns of EWS–FLI and FLI1 signal were detected (Supplemental Fig. S10). All motif associations were lost when the sites were permuted (Supplemental Fig. S11). These data suggest a large network of cooperative interactions for FLI1 binding, most frequently AP1 and GATA. Sites selective for EWS–FLI occupancy in sarcoma cells were distinguished in function, location, and composition from those sites that characterize endothelial targeting.

Chromatin features distinguish microsatellite repeats in Ewing Sarcoma

Since cell lineage dominated the variation in targeting of both chimeric and parental transcription factors, we explored features of chromatin configuration and histone modifications that could underlie these differences. We performed ChIP-seq in Ewing Sarcoma cells for histone marks associated with active (H3K4me1, H3K4me2, and H3K4me3) or repressed (H3K27me3) chromatin (Supplemental Tables 1, 6). We also performed Formaldehyde-Assisted Isolation of Regulatory Elements coupled with next-generation sequencing (FAIRE-seq) to identify regions of nucleosome depletion that characterize active regulatory elements (Supplemental Tables 1, 6). Consistent with other cell types, we found that both methylation on H3K4 and FAIRE demonstrated a strong association with gene expression, whereas H3K27me3 was inversely correlated with transcription (Fig. 5A).

Deregulation of homeobox genes is a common attribute of cancer (Cillo et al. 1999). We compared the histone methylation patterns across the four *HOX* clusters to those from embryonic stem cells and HUVEC (ENCODE Project Consortium 2004) (Supplemental Fig. S12A). Interestingly, at the *HOXA* cluster, we detected a bivalent signal similar to that observed in embryonic stem cells (Bernstein et al. 2006). At the other *HOX* clusters however, there was an overall lack of H3K27me3 and enrichment for H3K4me2 and H3K4me3. These marks associated with activation were not

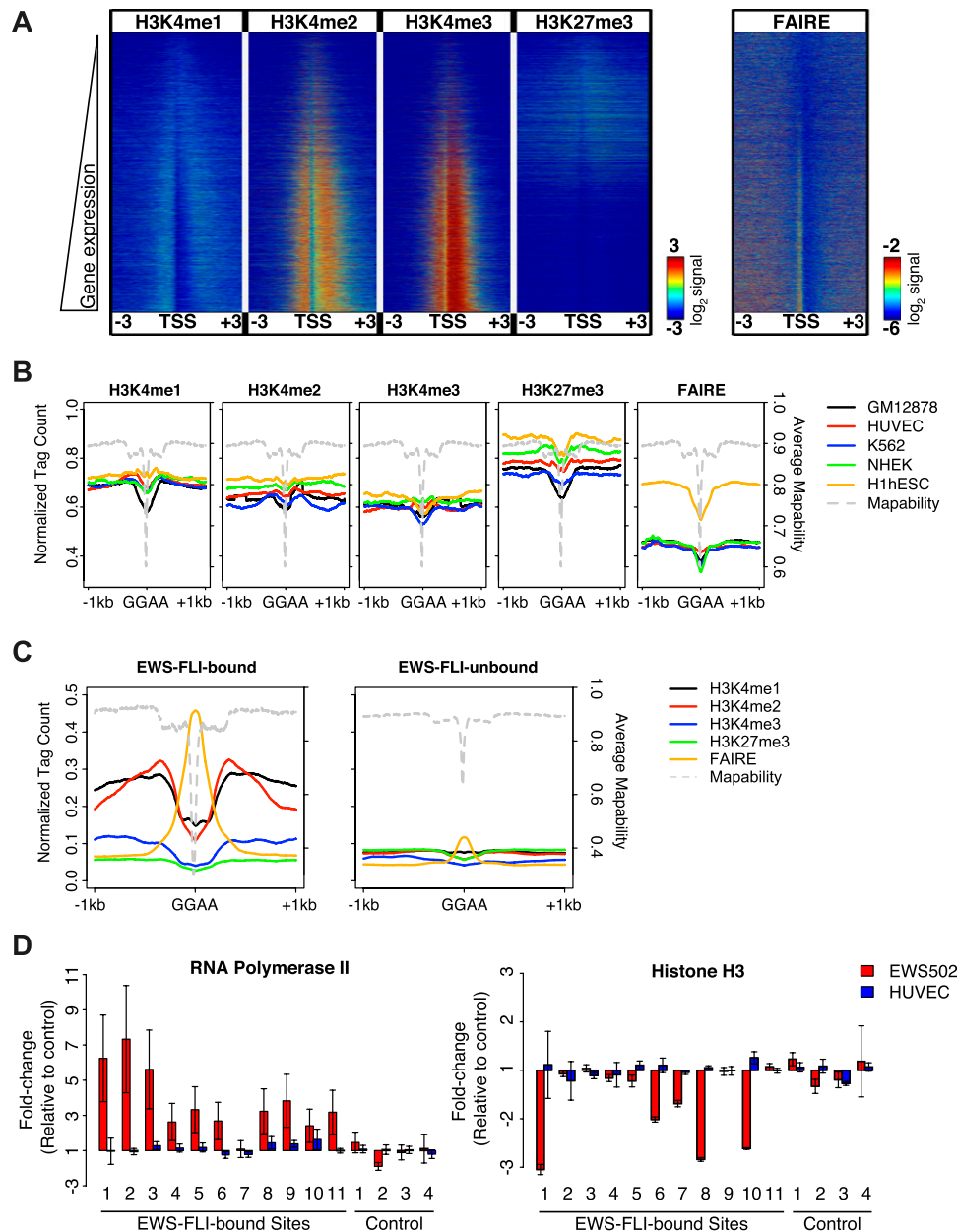


Figure 5. Deregulation of repetitive elements in Ewing Sarcoma. (A) Heatmap of normalized ChIP and FAIRE signal ± 3 kb around TSS ranked by gene expression in Ewing cells. Color was assigned on a \log_2 scale of -3 to 3 for ChIP and -6 to -2 for FAIRE. (B) Normalized ChIP and FAIRE signals around the centers of GGAA repeats in five ENCODE cell lines (GM12878, black; HUVEC, red; K562, blue; NHEK, green; H1hESC, orange). Mapability of the underlying DNA sequence is represented on a scale of 0 (ambiguous) to 1 (unique) and is plotted in gray. (C) Normalized ChIP and FAIRE signals around the centers of EWS-FLI-bound (left) or -unbound (right) GGAA repeats in Ewing Sarcoma cells. Mapability of the underlying DNA sequence is represented on a scale of 0 (ambiguous) to 1 (unique) and is plotted in gray. (D) Enrichment of EWS-FLI-bound GGAA repeats for RNA polymerase II (left) and histone H3 (right) in Ewing cells (red) and HUVEC (blue), as assayed by ChIP-qPCR. All values are represented as the fold-change relative to the average of the negative controls; fold-change values are centered on 1. Error bars represent the standard error from three technical replicates.

specific for a set of homeobox genes, in contrast to differentiated cells (e.g., HUVEC). Widespread activation across nearly all HOX genes in each cluster may contribute to the dedifferentiated or more stem-like state of Ewing Sarcomas. The loss of H3K27me3 signal was not observed genome-wide (e.g., *PAX2* and *WNT3A* loci) (Supplemental Fig. S12B).

We focused our analysis on regions containing the tetranucleotide repeats since they represented the most prominent feature

distinguishing EWS-FLI targeting. Comparing chromatin structure and histone methylation of the repeats in Ewing Sarcoma cells with HUVEC and other cell lines assayed as part of the ENCODE consortium, we observed that in virtually all cell types, repressive marks were common at GGAA microsatellites (Fig. 5B; Supplemental Fig. S13) and other repetitive elements (Mikkelsen et al. 2007; Meissner et al. 2008; Ernst and Kellis 2010). In Ewing cells, however, strong H3K4me1 and H3K4me2 signals flanked EWS-FLI-bound repeats

compared with those that were not bound, although the proportion of active histone marks that directly overlapped repeats was similar to other cell types (Fig. 5C; Supplemental Fig. S13). Moreover, EWS–FLI-bound sites were largely devoid of H3K27me3 and were nucleosome depleted. Relative to HUVEC and control regions, these sites were also bound by RNA polymerase II (Fig. 5D). To confirm nucleosome depletion, we also performed pan-histone H3 ChIP, which demonstrated overall histone H3 depletion at several sites (Fig. 5D). Intriguingly, other classes of repetitive elements were also enriched by FAIRE, including SINEs, LINEs, and other types of simple and microsatellite repeats (data not shown). Together, these data support the utilization of the microsatellites as transcriptional enhancers. We also observed a striking difference in DNA sequence-encoded nucleosome occupancy between EWS–FLI and FLI1-bound sites (Supplemental Fig. S14). Whereas FLI1-bound sequences demonstrated high nucleosome occupancy, a feature of regulatory elements in higher eukaryotes (Kaplan et al. 2009), this pattern was not seen for EWS–FLI-bound regions, further supporting the unique attributes of EWS–FLI targeted domains.

EWS–FLI targets enhancer like elements altering and maintaining the local chromatin environment

In addition to analyzing microsatellite regions we also assessed chromatin structure and histone modifications of each class of the differentially targeted regions identified by hierarchical clustering (Fig. 3A). These data further support that the chromatin around Ewing-specific clusters (clusters 1 and 2) exhibit a pattern resembling that of an enhancer element only in tumor cells (Fig. 6A). Conversely, the endothelial-specific clusters (clusters 4–6) show an

enhancer-like pattern unique to HUVEC. Distinct from the other clusters, sites in cluster 3, which exhibited binding by FLI1 in both cell types, were marked by H3K4me3. This chromatin signature supports our initial analyses (Fig. 2) that FLI1-specific sites tend to be located proximally to transcriptional start sites.

Since EWS–FLI-specific binding sites were normally nucleosome occupied in HUVEC and other cells, we asked whether EWS–FLI could alter nucleosome occupancy. EWS–FLI expression in HUVEC decreased nucleosome occupancy (indicated by increased FAIRE enrichment) at some of the closed chromatin regions, relative to control regions (Fig. 6B,C). Moreover, silencing of EWS–FLI in tumor cells resulted in increased nucleosome occupancy at all sites tested. These findings suggest that EWSR1 chimerism confers nucleosome displacement activity. Continued EWS–FLI expression is required for the maintenance of an open chromatin configuration at all sites tested.

Discussion

Lineage-specific outcomes are observed when chimeric transcription factors are expressed in various cell types, suggesting a major cell-specific influence on activity. One cell type may be permissive for transformation, whereas other cells may not tolerate expression, resulting in growth arrest or apoptosis. Cellular factors that influence activity may not be evident from studies of transcription factor targeting limited to a single transcriptional regulator in a single cell type.

Using an integrated genomic strategy to compare the oncoprotein EWS–FLI with its parental protein FLI1 in two relevant human cell types, we were able to separate the influence of chimerism and

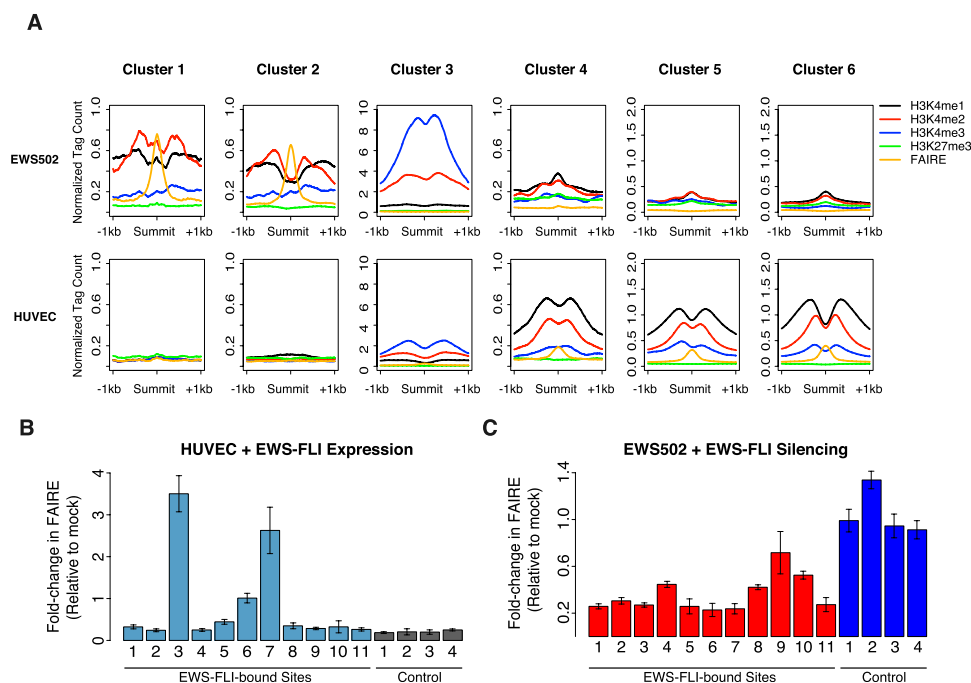


Figure 6. The EWS–FLI complex is capable of altering chromatin. (A) Normalized signals for H3K4me1 (black), H3K4me2 (red), H3K4me3 (blue), H3K27me3 (green), and FAIRE (orange) from both EWS502 and HUVEC are plotted for the 2-kb region surrounding the summits of sites identified by hierarchical clustering. (B) Change in FAIRE enrichment at EWS–FLI-bound GGAA repeats following EWS–FLI expression in HUVEC. All values are represented as fold-change relative to scrambled shRNA control. Error bars represent the standard error from three technical replicates. (C) Change in FAIRE enrichment at EWS–FLI-bound GGAA repeats following EWS–FLI silencing in EWS502. All values are represented as fold change relative to scrambled shRNA control. Error bars represent the standard error from three technical replicates.

cell type on genomic targeting and function. In tumor cells, chimerism resulted in genomic retargeting, with ~40% of EWS–FLI binding sites containing a tetranucleotide repeat composed of the core ETS motif. Although FLI1 can bind to these repeats, the majority of FLI1 sites contained the canonical ETS motif. In contrast, in endothelial cells, targeting of both proteins demonstrated remarkable similarity. EWS–FLI and FLI1 localized to sites containing the canonical ETS motif as well AP1 and GATA motifs. Binding to a number of other less-common DNA motifs suggests an extended network of interacting cooperative transcription factors. Given the abundance of ETS motif-containing sites in the genome, these interactions likely regulate cell lineage-specific patterns of genomic targeting. Although FLI1 can bind the tetranucleotide repeats both in vivo and in vitro, it fails to show activity at these repeats in reporter-based assays (Gangwal et al. 2008, 2010) suggesting that FLI1 requires the cooperation of other sequence-specific transcription factors to activate transcription. The association of ETS proteins with AP1 had been observed (Bassuk and Leiden 1995; Rao et al. 1999; Verger et al. 2001), and the functional association of EWS–FLI or FLI1 with AP1 and GATA1 has been demonstrated in cellular transformation and hematopoietic development (Verger et al. 2001; Kim et al. 2006; Stankiewicz and Crispino 2009). However, our data also suggest selectivity in cooperating transcription factors. Although studies of ETS1 identified cooperative binding with RUNX1 (Hollenhorst et al. 2009) and PAX5 (Garvie et al. 2001), neither relationship was evident in this study.

Differential targeting of EWS–FLI was influenced by chromatin structure and histone modifications. EWS–FLI-bound microsatellite regions in tumor cells that were atypically marked with an enhancer like signature bound RNA polymerase II, and resided in nucleosome-depleted regions. Our data suggest that through EWSR1 chimerism, the EWS–FLI complex gained nucleosome modification activity and is necessary to alter the local chromatin landscape, resulting in nucleosome depletion or destabilization. However, the observation of nucleosome depletion at repetitive regions suggests that other factors may initially create a favorable chromatin arrangement permitting EWS–FLI targeting, a question currently being explored. The presence of RNA polymerase II suggests that these regions may be transcribed, a feature recently shown to be common among human epithelial cancers (Ting et al. 2011).

This ability of EWS–FLI to alter chromatin structure is similar to that of FOXA1 or GATA4, which bind their cognate sites and affect chromatin configuration (Cirillo et al. 1998, 2002; Carroll et al. 2005, 2006; He et al. 2010; Hurtado et al. 2011). Since EWS–FLI does not contain the conserved motif thought to be required for core histone interactions, its activity may be mediated through the recruitment of histone modifying enzymes.

Comparative chromatin immunoprecipitation and the comprehensive identification of regulatory elements by FAIRE offer strategies to identify regions of the genome that might play a role in tumor development. One such example is the EWS–FLI-bound tetranucleotide repeat near *IGF1*, which exhibited length polymorphisms that may influence EWS–FLI targeting and gene regulation. A correlation between expression and the length of the repeat has been demonstrated for *NROB1* (Garcia-Aragoncillo et al. 2008). The identification of an extended tandem repeat proximal to *IGF1* may be of significance for disease development and treatment, since EWS–FLI-mediated IGF1 expression and signaling has been implicated in Ewing Sarcoma development (Scotlandi et al. 1996; Herrero-Martin et al. 2009) and inhibition of IGF1 signaling is being studied as a potential therapeutic strategy. Such polymorphisms could arise de novo during tumor development or represent an allelic

selection in individuals. The observed selection for longer repeats could represent a mechanism to augment target gene expression.

This study demonstrates the prospect of translational cancer genomics. Targeting the chromatin-modifying activity of the EWS–FLI complex could exploit the persistent “addiction” of the tumor to aberrant transcription and represents a novel therapeutic approach. Further work will be necessary to determine the functional significance of polymorphisms or other mutations on disease susceptibility, onset, progression, and treatment.

Methods

Cell culture

EWS502 were cultured in RPMI 1640 supplemented with 15% FBS and HUVEC cells were cultured in Vasculife Basal Media (Lifeline Technologies) supplemented with 10% FBS. Both cell lines were maintained at standard growth conditions of 37°C and 5% CO₂.

Lentiviral knockdown expression

A short hairpin region complementary to the 3′ untranslated region of FLI1 (5′-TGCCCATCCTGCACACTTACTTCAAGAGAGTAAGTGTGCAGGATGGGCTTTTC-3′ sense strand) together with PCR-generated HA-tagged EWS–FLI and HA-tagged FLI1 were cloned into pLL5.5 (Rubinson et al. 2003). Lentivirus was produced in HEK293T cells as described (Rubinson et al. 2003). EWS502 or HUVEC cells were infected with lentivirus in the presence of polybrene (6 µg/mL) for 3 h, after which the medium was changed. Chromatin or RNA was isolated at 72 h (see below).

Chromatin immunoprecipitation (ChIP) and Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE)

Chromatin was isolated and immunoprecipitation was performed as described in Davis et al. (2006) using 2 µg of anti-HA antibody (Abcam ab9110), anti-H3K4me1 (Abcam ab8895), anti-H3K4me2 (Abcam ab32356), anti-H3K4me3 (Abcam ab12209), anti-H3K27me3 (Millipore 07-449), anti-POLR2A (RNA pol II) (Abcam ab103968), or H3 (kindly provided by the Strahl lab). Immunoprecipitated DNA was prepared for high-throughput sequencing per the manufacturer's recommendations (Illumina), including DNA purification using AMPure XP beads (Agencourt) before PCR amplification. Quantitative PCR was performed as described (Absolute SYBR Green ROX Mix, Thermo Scientific). PCR primers are available upon request. FAIRE was performed on three independent cell harvests as previously described (Giresi et al. 2007), and isolated DNA was prepared for sequencing as above.

Quality control and reference genome alignment

Reads from chromatin immunoprecipitations were aligned to the reference human genome (hg18) with Bowtie (Langmead et al. 2009) using default parameters, and unambiguously placed reads were retained. Biological replicates were then merged, cross-replicate correlation was assessed, and reads were extended in silico to a final length of 200 bp. Any extended reads that overlapped large-scale repetitive elements were then removed. Reads from FAIRE were allowed to potentially map to up to four genomic locations, but the best scoring alignment was chosen. Biological replicates were then merged, cross-replicate correlation was assessed, and reads were extended in silico to a final length of 134 bp. Any extended reads that overlapped large-scale repetitive elements were then removed.

Peak calling, permutation, and repeat analysis

Areas of significant enrichment were identified using the Zero Inflated Negative Binomial Algorithm (Rashid et al. 2011). A window size and offset of 250 and 50 bp, respectively, were utilized for FLI1 and 500 and 125 bp for EWS-FLI. In all cases, a mixture regression model was created using a combination of the input control, GC content, and a background-derived copy-number variation model. Windows with *q*-values exceeding 0.95–0.99 were considered statistically significant, and peak boundaries were further refined when necessary. Additional parameters were specified to account for the broad nature of H3K27me3 domains. The percentage of peaks and average signal over a meta-gene were calculated using CEAS (Shin et al. 2009) and plotted in R. GGAA repeat regions were derived from RepeatMasker classes (GGAA)_n, (TTCC)_n, (GA)_n, (TC)_n, or other simple repeats that contained GGAA or its complement. For analyses of GGAA repeat length, peak coordinates were permuted 1000 times across the uniquely mappable genome while maintaining chromosomal distribution using BEDTools (Quinlan and Hall 2010). The frequency of tandem GGAA/TTCC repeats was computed for lengths 1–25 and compared with that of the test peak coordinates to compute a two-sided *P*-value.

Hierarchical clustering and motif identification

A union set of all EWS-FLI and FLI1 peaks for each cell type were merged using Galaxy (Goecks et al. 2010). For each of the 51,085 regions, we retrieved the average number of sequencing reads from each experiment and filtered for regions where the standard deviation and interquartile range exceeded 0.75 and 0.5, respectively. The resulting 6525 regions were median centered and hierarchically clustered using average linkage and Pearson correlation. The resulting dendrogram and heatmap were created in Java Treeview (Saldanha 2004). Regions identified by clustering analysis were narrowed to a refined window immediately around the location of binding by intersecting the union set of all 200-bp windows around the site of greatest signal (summit). De novo motif detection was performed with CisFinder (Sharov and Ko 2009) using the 200-bp flanking sequence as background. Motif heatmaps were created by calculating the input-normalized number of sequencing reads for each sample in the 2-kb region surrounding each identified motif location.

Microarray analysis and gene permutation

RNA was extracted from $\sim 1.5 \times 10^7$ lentivirally infected EWS502 and HUVEC cells (RNeasy, Qiagen). After quality assessment (Bioanalyzer, Agilent) RNA was hybridized to Affymetrix GeneChip Human Exon 1.0 ST arrays by the UNC Microarray Facility. Differentially expressed genes were determined by an ANOVA (Partek Genomics Suite, *P* < 0.05) comparing gene expression under conditions of EWS-FLI or FLI1 knockdown only to knockdown with ectopic transcription factor expression. Only genes with a RefSeq identification number that mapped to standard chromosomes were considered. Differentially expressed genes were assessed for significantly enriched biological pathways and functions using Ingenuity Pathway Analysis (*P* < 0.05). To compare the number of differentially expressed genes close to binding sites as identified by hierarchical clustering to an expected distribution, we permuted over all RefSeq genes 10,000 times and computed a two-sided *P*-value accordingly.

Data access

All sequencing and microarray data have been submitted to the NCBI Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE31838.

Acknowledgments

We acknowledge the input of the Davis, Lieb, Strahl, and Rathmell labs, and in particular the contributions of the following individuals: Paul Giresi for constructive advice and assistance with FAIRE; James Bear for assistance with gene silencing; Wei Sun, Joel Parker, and Naim Rashid for statistical support; Piotr Mieczkowski, Hemant Kelkar, and the UNC High Throughput Sequencing Facility; and the UNC Microarray Core facility. We gratefully acknowledge support from the NCI/NIH (K08CA100400), V Foundation for Cancer Research, the Rita Allen Foundation, and the Corn-Hammond Fund for Pediatric Oncology.

Authors' contributions: The work presented here was carried out in collaboration between all authors. I.J.D. conceived and designed the approach. M.P., J.M.S., M.D.I., S.B.W., and A.W.M. designed methods and performed the experiments. M.P., J.M.S., J.D.L., and I.J.D. interpreted the results and wrote the manuscript. All authors have contributed to, seen, and approved of the manuscript.

References

- Bartel FO, Higuchi T, Spyropoulos DD. 2000. Mouse models in the study of the Ets family of transcription factors. *Oncogene* **19**: 6443–6454.
- Bassuk AG, Leiden JM. 1995. A direct physical association between ETS and AP-1 transcription factors in normal human T cells. *Immunity* **3**: 223–237.
- Benjamin LE, Fredericks WJ, Barr FG, Rauscher FJ 3rd. 1996. Fusion of the EWS1 and WT1 genes as a result of the t(11;22)(p13;q12) translocation in desmoplastic small round cell tumors. *Med Pediatr Oncol* **27**: 434–439.
- Bernheim A. 2010. Cytogenomics of cancers: from chromosome to sequence. *Mol Oncol* **4**: 309–322.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**: 315–326.
- Bertolotti A, Melot T, Acker J, Vigneron M, Delattre O, Tora L. 1998. EWS, but not EWS-FLI-1, is associated with both TFIIID and RNA polymerase II: interactions between two members of the TET family, EWS and hTAFII68, and subunits of TFIIID and RNA polymerase II complexes. *Mol Cell Biol* **18**: 1489–1497.
- Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoutte J, Shao W, Hestermann EV, Geistlinger TR, et al. 2005. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**: 33–43.
- Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoutte J, Brodsky AS, Keeton EK, Fertuck KC, Hall GF, et al. 2006. Genome-wide analysis of estrogen receptor binding sites. *Nat Genet* **38**: 1289–1297.
- Cillo C, Faiella A, Cantile M, Boncinelli E. 1999. Homeobox genes and cancer. *Exp Cell Res* **248**: 1–9.
- Cirillo LA, McPherson CE, Bossard P, Stevens K, Cherian S, Shim EY, Clark KL, Burley SK, Zaret KS. 1998. Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO J* **17**: 244–254.
- Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol Cell* **9**: 279–289.
- Davis IJ, Kim JJ, Ozsolak F, Widlund HR, Rozenblatt-Rosen O, Granter SR, Du J, Fletcher JA, Denny CT, Lessnick SL, et al. 2006. Oncogenic MITF dysregulation in clear cell sarcoma: defining the Mit family of human cancers. *Cancer Cell* **9**: 473–484.
- Delattre O, Zucman J, Plougastel B, Desmaziere C, Melot T, Peter M, Kovar H, Joubert I, de Jong P, Rouleau G, et al. 1992. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359**: 162–165.
- Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwierz P, Alekseyenko AA, Cheung MS, Day DS, Gadel S, Gorchakov AA, et al. 2010. An assessment of histone-modification antibody quality. *Nat Struct Mol Biol* **18**: 91–93.
- ENCODE Project Consortium. 2004. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* **306**: 636–640.
- Ernst J, Kellis M. 2010. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol* **28**: 817–825.
- Gangwal K, Lessnick SL. 2008. Microsatellites are EWS/FLI response elements: genomic “junk” is EWS/FLI's treasure. *Cell Cycle* **7**: 3127–3132.
- Gangwal K, Sankar S, Hollenhorst PC, Kinsey M, Haroldsen SC, Shah AA, Boucher KM, Watkins WS, Jorde LB, Graves BJ, et al. 2008.

- Microsatellites as EWS/FLI response elements in Ewing's sarcoma. *Proc Natl Acad Sci* **105**: 10149–10154.
- Gangwal K, Close D, Enriquez CA, Hill CP, Lessnick SL. 2010. Emergent properties of EWS/FLI regulation via GGAA microsatellites in ewing's sarcoma. *Genes Cancer* **1**: 177–187.
- Garcia-Aragoncillo E, Carrillo J, Lalli E, Agra N, Gomez-Lopez G, Pestana A, Alonso J. 2008. DAX1, a direct target of EWS/FLI1 oncoprotein, is a principal regulator of cell-cycle progression in Ewing's tumor cells. *Oncogene* **27**: 6034–6043.
- Garvie CW, Hagman J, Wolberger C. 2001. Structural studies of Ets-1/Pax5 complex formation on DNA. *Mol Cell* **8**: 1267–1276.
- Giresi PG, Kim J, McDaniel RM, Iyer VR, Lieb JD. 2007. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res* **17**: 877–885.
- Goecks J, Nekrutenko A, Taylor J. 2010. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* **11**: R86. doi: 10.1186/gb-201011-8-r86.
- Guillon N, Tirode F, Boeva V, Zynovyev A, Barillot E, Delattre O. 2009. The oncogenic EWS-FLI1 protein binds in vivo GGAA microsatellite sequences with potential transcriptional activation function. *PLoS ONE* **4**: e4932. doi: 10.1371/journal.pone.0004932.
- He HH, Meyer CA, Shin H, Bailey ST, Wei G, Wang Q, Zhang Y, Xu K, Ni M, Lupien M, et al. 2010. Nucleosome dynamics define transcriptional enhancers. *Nat Genet* **42**: 343–347.
- Herrero-Martin D, Osuna D, Ordóñez JL, Sevillano V, Martins AS, Mackintosh C, Campos M, Madoz-Gurpide J, Otero-Motta AP, Caballero G, et al. 2009. Stable interference of EWS-FLI1 in an Ewing sarcoma cell line impairs IGF-1/IGF-1R signalling and reveals TOPK as a new target. *Br J Cancer* **101**: 80–90.
- Hollenhorst PC, Chandler KJ, Poulsen RL, Johnson WE, Speck NA, Graves BJ. 2009. DNA specificity determinants associate with distinct transcription factor functions. *PLoS Genet* **5**: e1000778. doi: 10.1371/journal.pgen.1000778.
- Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS. 2011. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat Genet* **43**: 27–33.
- Jaishankar S, Zhang J, Roussel MF, Baker SJ. 1999. Transforming activity of EWS/FLI1 is not strictly dependent upon DNA-binding activity. *Oncogene* **18**: 5592–5597.
- Jeon IS, Davis JN, Braun BS, Sublett JE, Roussel MF, Denny CT, Shapiro DN. 1995. A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1. *Oncogene* **10**: 1229–1234.
- Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J, et al. 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**: 362–366.
- Kauer M, Ban J, Kofler R, Walker B, Davis S, Meltzer P, Kovar H. 2009. A molecular function map of Ewing's sarcoma. *PLoS ONE* **4**: e5415. doi: 10.1371/journal.pone.0005415.
- Kim S, Denny CT, Wisdom R. 2006. Cooperative DNA binding with AP-1 proteins is required for transformation by EWS-Ets fusion proteins. *Mol Cell Biol* **26**: 2467–2478.
- Kinsey M, Smith R, Lessnick SL. 2006. NR0B1 is required for the oncogenic phenotype mediated by EWS/FLI1 in Ewing's sarcoma. *Mol Cancer Res* **4**: 851–859.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25. doi: 10.1186/gb-2009-10-3-r25.
- Lessnick SL, Braun BS, Denny CT, May WA. 1995. Multiple domains mediate transformation by the Ewing's sarcoma EWS/FLI-1 fusion gene. *Oncogene* **10**: 423–431.
- Lessnick SL, Dacwag CS, Golub TR. 2002. The Ewing's sarcoma oncoprotein EWS/FLI1 induces a p53-dependent growth arrest in primary human fibroblasts. *Cancer Cell* **1**: 393–401.
- Li H, Watford W, Li C, Parmelee A, Bryant MA, Deng C, O'Shea J, Lee SB. 2007. Ewing sarcoma gene EWS is essential for meiosis and B lymphocyte development. *J Clin Invest* **117**: 1314–1323.
- Liu F, Walmsley M, Rodaway A, Patient R. 2008. FLI1 acts at the top of the transcriptional network driving blood and endothelial development. *Curr Biol* **18**: 1234–1240.
- Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, et al. 2003. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* **31**: 374–378.
- May WA, Gishizky ML, Lessnick SL, Lunsford LB, Lewis BC, Delattre O, Zucman J, Thomas G, Denny CT. 1993a. Ewing sarcoma 11;22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by FLI1 for transformation. *Proc Natl Acad Sci* **90**: 5752–5756.
- May WA, Lessnick SL, Braun BS, Klemsz M, Lewis BC, Lunsford LB, Hromas R, Denny CT. 1993b. The Ewing's sarcoma EWS/FLI-1 fusion gene encodes a more potent transcriptional activator and is a more powerful transforming gene than FLI-1. *Mol Cell Biol* **13**: 7393–7398.
- McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**: 495–501.
- Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, et al. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**: 766–770.
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**: 553–560.
- Miyagawa Y, Okita H, Nakajima H, Horiuchi Y, Sato B, Taguchi T, Toyoda M, Katagiri YU, Fujimoto J, Hata J, et al. 2008. Inducible expression of chimeric EWS/ETS proteins confers Ewing's family tumor-like phenotypes to human mesenchymal progenitor cells. *Mol Cell Biol* **28**: 2125–2137.
- Ohno T, Ouchida M, Lee L, Gatalica Z, Rao VN, Reddy ES. 1994. The EWS gene, involved in Ewing family of tumors, malignant melanoma of soft parts and desmoplastic small round cell tumors, codes for an RNA binding protein with novel regulatory domains. *Oncogene* **9**: 3087–3097.
- Pustazzeri MP, Seelentag W, Bosman FT. 2006. Immunohistochemical expression of endothelial markers CD31, CD34, von Willebrand factor, and FLI-1 in normal human tissues. *J Histochem Cytochem* **54**: 385–395.
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.
- Rao VN, Ohno T, Prasad DD, Bhattacharya G, Reddy ES. 1993. Analysis of the DNA-binding and transcriptional activation functions of human FLI-1 protein. *Oncogene* **8**: 2167–2173.
- Rao S, Matsumura A, Yoon J, Simon MC. 1999. SPI-B activates transcription via a unique proline, serine, and threonine domain and exhibits DNA binding affinity differences from PU.1. *J Biol Chem* **274**: 11115–11124.
- Rashid N, Giresi PG, Ibrahim JG, Sun W, Lieb JD. 2011. ZINBA integrates local covariates with DNA-seq data to identify broad and narrow regions of enrichment, even within amplified genomic regions. *Genome Biol* **12**: R67. doi: 10.1186/gb-2011-12-7-r67.
- Riggi N, Cironi L, Provero P, Suva ML, Kaloulis K, Garcia-Echeverria C, Hoffmann F, Trumpp A, Stamenkovic I. 2005. Development of Ewing's sarcoma from primary bone marrow-derived mesenchymal progenitor cells. *Cancer Res* **65**: 11459–11468.
- Riggi N, Suva ML, Suva D, Cironi L, Provero P, Percier S, Joseph JM, Stehle JC, Baumer K, Kindler V, et al. 2008. EWS-FLI-1 expression triggers a Ewing's sarcoma initiation program in primary human mesenchymal stem cells. *Cancer Res* **68**: 2176–2185.
- Rubinson DA, Dillon CP, Kwiatkowski AV, Sievers C, Yang L, Kopinja J, Rooney DL, Zhang M, Ihrig MM, McManus MT, et al. 2003. A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. *Nat Genet* **33**: 401–406.
- Saldanha AJ. 2004. Java Treeview—extensible visualization of microarray data. *Bioinformatics* **20**: 3246–3248.
- Scotlandi K, Benini S, Sarti M, Serra M, Lollini PL, Maurici D, Picci P, Manara MC, Baldini N. 1996. Insulin-like growth factor I receptor-mediated circuit in Ewing's sarcoma/peripheral neuroectodermal tumor: a possible therapeutic target. *Cancer Res* **56**: 4570–4574.
- Sharov AA, Ko MS. 2009. Exhaustive search for over-represented DNA sequence motifs with CisFinder. *DNA Res* **16**: 261–273.
- Shin H, Liu T, Manrai AK, Liu XS. 2009. CEAS: cis-regulatory element annotation system. *Bioinformatics* **25**: 2605–2606.
- Shing DC, McMullan DJ, Roberts P, Smith K, Chin SF, Nicholson J, Tillman RM, Ramani P, Cullinane C, Coleman N. 2003. FUS/ERG gene fusions in Ewing's tumors. *Cancer Res* **63**: 4568–4576.
- Smith R, Owen LA, Trem DJ, Wong JS, Whangbo JS, Golub TR, Lessnick SL. 2006. Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma. *Cancer Cell* **9**: 405–416.
- Sorensen PH, Lessnick SL, Lopez-Terrada D, Liu XF, Triche TJ, Denny CT. 1994. A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nat Genet* **6**: 146–151.
- Spyropoulos DD, Pharr PN, Lavenburg KR, Jackers P, Papas TS, Ogawa M, Watson DK. 2000. Hemorrhage, impaired hematopoiesis, and lethality in mouse embryos carrying a targeted disruption of the FLI1 transcription factor. *Mol Cell Biol* **20**: 5643–5652.
- Stankiewicz MJ, Crispino JD. 2009. ETS2 and ERG promote megakaryopoiesis and synergize with alterations in GATA-1 to immortalize hematopoietic progenitor cells. *Blood* **113**: 3337–3347.

- Tanaka K, Iwakuma T, Harimaya K, Sato H, Iwamoto Y. 1997. EWS-Flt1 antisense oligodeoxynucleotide inhibits proliferation of human Ewing's sarcoma and primitive neuroectodermal tumor cells. *J Clin Invest* **99**: 239–247.
- Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, Coffman EJ, Contino G, Deshpande V, Iafrate AJ, Letovsky S, et al. 2011. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* **331**: 593–596.
- Tomkins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, et al. 2005. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **310**: 644–648.
- Verger A, Buisine E, Carrere S, Wintjens R, Flourens A, Coll J, Stehelin D, Dutourque-Coquillaud M. 2001. Identification of amino acid residues in the ETS transcription factor Erg that mediate Erg-Jun/Fos-DNA ternary complex formation. *J Biol Chem* **276**: 17181–17189.
- Wei GH, Badis G, Berger ME, Kivioja T, Palin K, Enge M, Bonke M, Jolma A, Varjosalo M, Gehrke AR, et al. 2010. Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J* **29**: 2147–2160.
- Xie X, Rigor P, Baldi P. 2009. MotifMap: a human genome-wide map of candidate regulatory motif sites. *Bioinformatics* **25**: 167–174.
- Zucman J, Delattre O, Desmaziere C, Epstein AL, Stenman G, Speleman F, Fletcher CD, Aurias A, Thomas G. 1993. EWS and ATF-1 gene fusion induced by t(12;22) translocation in malignant melanoma of soft parts. *Nat Genet* **4**: 341–345.

Received May 1, 2011; accepted in revised form September 13, 2011.